DOCUMENT RESUME

ED 184 914 SO 012 383

AUTHOR Richard, James M., Jr.
TITLE Regression Weights and Communication among Researchers from Different Disciplines.
INSTITUTION Johns Hopkins Univ., Baltimore, Md. Center for Social Organization of Schools.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO CSOS-R-283
PUB DATE Sep 79
GRANT NIE-G-78-0210
NOTE 37p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Comparative Analysis; Data Analysis; *Differences; Error of Measurement; *Multiple Regression Analysis; Psychology; Raw Scores; *Research Methodology; *Social Science Research; Sociology; *Statistical Analysis

ABSTRACT

This report examines the difference in approaches between sociologists and psychologists when using multiple regression techniques in the analysis of behavioral data. Psychologists and sociologists are often divided in their orientation toward regression techniques, and this division could be a substantial and unfortunate barrier to communication among researchers. Many sociological methodologists argue that unstandardized (raw score) regression weights should be used in comparing groups, whereas psychologists rely almost exclusively on correlations and standard score regression weights. This report attempts to clarify the issue by illustrating how raw and standard score regression weights behave in various circumstances. The report is presented in four sections. Section one explores the reasons why sociologists prefer unstandardized weights. Section two discusses reasons why psychologists prefer standardized weights. Section three examines circumstances that do not fit conveniently under the rubric of comparing sociological and psychological perspectives on regression weights. The concluding section considers the implications of these analyses. It is found that each kind of weight is invariant in some circumstances where the other is not, and each kind can provide scientifically meaningful information that the other cannot. (Author/AV)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED184914

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
J. Hollifield
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Regression Weights and Communication Among Researchers From Different Disciplines

Grant No. NIE-G-78-0210

SP 012 383

James M. Richards, Jr.

Report No. 283

September, 1979

Published by the Center for Social Organization of Schools, supported in part as a research and development center by funds from the United States National Institute of Education, Department of Health, Education and Welfare. The opinions expressed in this publication do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by the Institute should be inferred.

The Johns Hopkins University

Baltimore, Maryland

MAR 7 1980

Introductory Statement

The Center for Social Organization of Schools has two primary objectives: to develop a scientific knowledge of how schools affect their students, and to use this knowledge to develop better school practices and organization.

The Center works through four programs to achieve its objectives. The Policy Studies in School Desegregation program applies the basic theories of social organization of schools to study the internal conditions of desegregated schools, the feasibility of alternative desegregation policies, and the interrelation of school desegregation with other equity issues such as housing and job desegregation. The School Organization program is currently concerned with authority-control structures, task structures, reward systems, and peer group processes in schools. It has produced a large-scale study of the effects of open schools, has developed Student Team Learning Instructional processes for teaching various subjects in elementary and secondary schools, and has produced a computerized system for school-wide attendance monitoring. The School Process and Career Development program is studying transitions from high school to post secondary institutions and the role of schooling in the development of career plans and the actualization of labor market outcomes. The Studies in Delinquency and School Environments program is examining the interaction of school environments, school experiences, and individual characteristics in relation to in-school and later-life delinquency.

This report, prepared by the School Organization program, compares unstandardized, or raw score, regression weights with standardized regression weights. The concrete illustrations of how these weights behave in various specific circumstances should help educational and other researchers choose which type of regression weight, if any, to use in analyzing their data.

Regression Weights and Communication Among

Researchers from Different Disciplines

Abstract

Different orientations about multiple regression techniques could be a substantial barrier to communication among psychologists and sociologists. Most recent sociological methodologists argue strongly that only raw score regression weights should be used in comparing groups, while psychologists rely almost exclusively on correlations and standard score regression weights. This review attempted to clarify this issue by illustrating how raw and standard score regression weights behave in various circumstances. These illustrations indicate that each kind of weight is invariant in some circumstances where the other is not, and each kind can provide scientifically meaningful information that the other cannot. On the other hand, most of the illustrated circumstances affect both kinds of regression weights, sometimes quite strongly. Thus neither kind of weight appears particularly robust, and certainly neither kind appears to provide a panacea for overcoming the problems of analyzing behavioral data.

Regression Weights and Communication Among Researchers from Different Disciplines

Psychologists and sociologists often are divided in their orientations toward multiple correlation and regression techniques, and this division could be a substantial barrier to communication among researchers. The use of unstandardized (i.e., raw score), or "metric," regression weights has become a substantial issue among sociologists, with most recent methodologists, especially those influenced by econometricians, arguing strongly that only such weights should be used in comparing groups (Blalock, 1964, 1967, 1968, 1976; Bohrnstedt, 1969; Duncan, 1975; Kim & Mueller, 1976; Schoenberg, 1972). On the other hand, most psychologists continue to base their analyses on correlations and standardized (i.e., standard score) regression weights, and appear largely unaware of, or uninterested in, the entire issue. For example, the recent controversy among psychologists about whether random or equal regression weights are as accurate as best fitting weights (Dawes & Corrigan, 1974; Laughlin, 1978; Pruzek & Frederick, 1978; Wainer, 1976, 1978) makes sense only if the argument is limited to standardized weights. Yet none of the participants in the controversy considered this limitation important enough to discuss in detail.

Researchers from different disciplines should be able to communicate with each other, and this particular division appears especially unfortunate because regression techniques are central to such important procedures as path analysis (Cooley, 1978; Duncan, 1966; Werts & Linn, 1970), structural equation modeling (Duncan, 1975), and the analysis of variance-covariance structures (Jöreskog, 1973). Accordingly, the present review was undertaken to help clarify the standardized vs. unstandardized

weights issue by presenting some concrete illustrations of how these weights behave in various circumstances. Although the relationships that are reviewed are well known (Gulliksen, 1950), in the past they rarely if ever have been used to compare standardized and unstandardized regression weights. Moreover, the few sources that have discussed both kinds of weights, such as the SPSS manual (Nie, Hull, Jenkins, Steinbrenner, & Brent, 1975), are of little help in understanding, much less resolving, the differences between the sociological and psychological perspectives.

This review is divided into four sections. The first section explores the reasons why sociologists prefer unstandardized weights. The second examines the reasons why psychologists prefer standardized weights. A third examines circumstances that do not fit conveniently under the rubric of comparing sociological and psychological perspectives on regression weights. The concluding section considers the implications of these analyses.

Why Do Sociologists Prefer Unstandardized Weights?

Bohrnstedt (1969) provided the following representative example of the major reasons sociologists cite for preferring unstandardized regression weights:

> "The regression coefficient is suggested as more appropriate than the correlation coefficient or the standardized coefficient since the unstandardized coefficient is relatively stable across subsamples of a population whereas the standardized coefficients...

may vary significantly as a function of the standard deviations."

What sociologists mean by such statements can be clarified by considering the familiar example of college admissions. Studying the relationship between high school and college grades (GPAs) is problematical because students are selected largely on the basis of high school grades. In the pure case, students falling below some minimum grade are denied admission to college, thus truncating the distribution of high school grades. Such truncation directly reduces the standard deviation of high school grades and indirectly reduces both the correlation between high school and college GPAs and the standard deviation of college grades (Gulliksen, 1950). These effects would be substantially greater at more selective colleges, so such truncation is an extreme case of the kind of variation in standard deviations cited by Bohrnstedt (1969).

Accordingly, formulas presented by Gulliksen (1950, pp. 128-144) were used to compute the effect of varying degrees of selection on hypothetical relationships between high school and college grades, and especially the impact of selection on the standardized and unstandardized regression weights. It is customary to call the variables on which cases are selected (in this example, high school grades) the "explicit selection" variables and to call the variables that are indirectly affected by selection (in this example, college grades) the "incidental selection" variables. The formulas require knowledge of the standard deviation of the selected group on the explicit selection variable, so these standard deviations were computed from the table of randomly distributed cases presented by Ghiselli (1969). Results are summarized

in Table 1. The data in the left column of this table are the underlying true or population values. Although these data are hypothetical, they were chosen to be close to the values suggested by compendia (American College Testing Program, 1966; Lavin, 1965). It should be emphasized that the results in Table 1 would occur whether or not high school grades "cause" college grades.

Table 1

Effects of Various Levels of Selection on the Independent Variable of High School Grades

	Selected Group Includes Top:			
	100%	80%	50%	20%
High School GPA				
Mean	2.60	2.84	3.16	3.58
S.D.	.70	.53	.42	.32
College GPA				
Mean	2.10	2.28	2.52	2.83
S.D.	.80	.73	.68	.65
Correlation Between GPAs	.65	.55	.46	.36
Regression Weights				
Standardized	.65	.55	.46	.36
Unstandardized	.74	.74	.74	.74

Note--In deriving this and subsequent tables, computations were carried out to six decimal places then rounded to two.

The most important trends in this table are that selection on high school grades reduced the correlation between high school and college grades and therefore the standardized regression weights, while the unstandardized regression weights were invariant under such selection. This invariance is just the property of unstandardized weights which was cited by Bohrnstedt (1969) and which underlies most of sociologists' preference for them. Although Gulliksen (1950) mentions this property of unstandardized weights, like other psychologists he does not appear to consider it very important.

The results in Table 1 describe an absolute relationship between high school and college grades in which the only effect of selection is to eliminate cases. In other words, Table 1 describes a situation where students who attend highly selective colleges receive exactly the same grades they would have received if those colleges were completely unselective. There should be no confusion about this point; much of sociologists' preference for unstandardized weights is based on a search for, and an expectation of finding, relationships that are invariant in precisely this absolute sense. When such absolute relationships in fact exist, unstandardized weights do offer notable advantages. To illustrate these advantages, consider the following example suggested by Kim and Mueller (1976).

Suppose that several companies pursue identical sexually and racially discriminatory policies so that the rate of pay for their workers is determined by the formula:

$$\text{Pay Rate} = \$10 + \$2 \times \text{Sex (Male=1, Female=0)} + \$3 \times \text{Race (White=1, Other=0)} + \$1 \times \text{years of experience} \quad (1)$$

Table 2 then shows how variations among companies in the sexual composition, racial composition, and experience of their work forces affect the standardized and unstandardized regression weights relating those variables to pay rate. For computational convenience this table assumes no correlation among sex, race, and experience, but this assumption has no effect on the overall pattern of results (Kim & Mueller, 1976). The most important characteristic of this example is that no errors of specification or measurement are present. (Specification error refers, roughly, to including an irrelevant variable or failing to include a relevant variable in a regression study of causal relationships.)

Table 2

Standardized and Unstandardized Regression Weights
for Hypothetical Example Suggested by Kim and Mueller

% Male	% White	Range in Years of Experience	Sex		Race		Experience		Pay		Standardized Regression Weight for			Unstandardized Regression Weight for		
			Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Sex	Race	Exp	Sex	Race	Exp
20%	20%	0-6	.20	.40	.20	.40	3.00	1.00	14.00	1.76	.46	.68	.57	2.00	3.00	1.00
		0-12	.20	.40	.20	.40	6.00	2.00	17.00	2.47	.32	.49	.81	2.00	3.00	1.00
		0-18	.20	.40	.20	.40	9.00	3.00	20.00	3.33	.24	.36	.90	2.00	3.00	1.00
	50%	0-6	.20	.40	.50	.50	3.00	1.00	14.90	1.97	.41	.76	.51	2.00	3.00	1.00
		0-12	.20	.40	.50	.50	6.00	2.00	17.90	2.62	.30	.57	.76	2.00	3.00	1.00
		0-18	.20	.40	.50	.50	9.00	3.00	20.90	3.45	.23	.44	.87	2.00	3.00	1.00
	80%	0-6	.20	.40	.80	.40	3.00	1.00	15.80	1.76	.46	.68	.57	2.00	3.00	1.00
		0-12	.20	.40	.80	.40	6.00	2.00	18.80	2.47	.32	.49	.81	2.00	3.00	1.00
		0-18	.20	.40	.80	.40	9.00	3.00	21.80	3.33	.24	.36	.90	2.00	3.00	1.00
50%	20%	0-6	.50	.50	.20	.40	3.00	1.00	14.60	1.85	.54	.65	.54	2.00	3.00	1.00
		0-12	.50	.50	.20	.40	6.00	2.00	17.60	2.54	.39	.47	.79	2.00	3.00	1.00
		0-18	.50	.50	.20	.40	9.00	3.00	20.60	3.38	.30	.35	.89	2.00	3.00	1.00
	50%	0-6	.50	.50	.50	.50	3.00	1.00	15.50	2.06	.49	.73	.49	2.00	3.00	1.00
		0-12	.50	.50	.50	.50	6.00	2.00	18.50	2.69	.37	.56	.74	2.00	3.00	1.00
		0-18	.50	.50	.50	.50	9.00	3.00	21.50	3.50	.29	.43	.86	2.00	3.00	1.00
	80%	0-6	.50	.50	.80	.40	3.00	1.00	16.40	1.85	.54	.65	.54	2.00	3.00	1.00
		0-12	.50	.50	.80	.40	6.00	2.00	19.40	2.54	.39	.47	.79	2.00	3.00	1.00
		0-18	.50	.50	.80	.40	9.00	3.00	22.40	3.38	.30	.35	.89	2.00	3.00	1.00
80%	20%	0-6	.80	.40	.20	.40	3.00	1.00	15.20	1.76	.46	.68	.57	2.00	3.00	1.00
		0-12	.80	.40	.20	.40	6.00	2.00	18.20	2.47	.32	.49	.81	2.00	3.00	1.00
		0-18	.80	.40	.20	.40	9.00	3.00	21.20	3.33	.24	.36	.90	2.00	3.00	1.00
	50%	0-6	.80	.40	.50	.50	3.00	1.00	16.10	1.97	.41	.76	.51	2.00	3.00	1.00
		0-12	.80	.40	.50	.50	6.00	2.00	19.10	2.62	.30	.57	.76	2.00	3.00	1.00
		0-18	.80	.40	.50	.50	9.00	3.00	22.10	3.45	.23	.44	.87	2.00	3.00	1.00
	80%	0-6	.80	.40	.80	.40	3.00	1.00	17.00	1.76	.46	.68	.57	2.00	3.00	1.00
		0-12	.80	.40	.80	.40	6.00	2.00	20.00	2.47	.32	.49	.81	2.00	3.00	1.00
		0-18	.80	.40	.80	.40	9.00	3.00	23.00	3.33	.24	.36	.90	2.00	3.00	1.00

Note-It is assumed that sex, race and experience are uncorrelated so the standardized regression weights are equal to the zero-order correlations.

The pattern of results is clear. The standardized regression weights (which in this case are equal to the zero-order correlations) for each variable varied not only with changes in its own standard deviation but also with changes in the standard deviations of the other two causal variables, findings that help explain why some writers argue that the percent of variance accounted for by various variables may be of little scientific interest (Duncan, 1975). The unstandardized regression weights, on the other hand, were invariant and in addition clearly revealed the underlying causal relationships.

A final, less important reason for sociologists' preference for unstandardized weights involves the impact of error variance on the regression weights (Bohrnstedt & Carter, 1971). Table 3 compares the effects of independent variable error and dependent variable error.

Table 3

Impact of Measurement
Error on Regression Weights

	S.D. of Error Scores			
	0.00	2.50	5.00	10.00
Error-Free Variable				
Mean	150.00	150.00	150.00	150.00
S.D.	20.00	20.00	20.00	20.00
Error-Prone Variable				
Mean	150.00	150.00	150.00	150.00
S.D.	10.00	10.31	11.81	14.14
Correlation Between Variables	.80	.78	.72	.57
Regression Weights				
Standardized	.80	.78	.72	.57
Unstandardized for Error-Prone				
Independent Variable	1.60	1.50	1.28	.80
Dependent Variable	.40	.40	.40	.40

Again the pattern is clear. The correlation and therefore the standardized regression weights were reduced by error variance on either the independent

variable or the dependent variable, while the unstandardized weights varied with independent variable error but were invariant for dependent variable error.

Why Do Psychologists Prefer Standardized Weights?

Perhaps the most basic reason for psychologists' preference for standardized weights is that they believe that most scales in the behavioral sciences measure how people stand relative to some norm rather than their absolute standing on some universal quality. For example, psychologists are likely to believe that the grades which college students receive depend on how they compare with other students currently attending their college, so in general students who attend highly selective colleges will not receive exactly the same grade they would have received if those colleges were completely unselective. In other words, psychologists doubt the existence in the real world of the kind of absolute relationships that underlie the argument for unstandardized weights, and of the kind of data that would be prevalent if the argument were correct (Linn, 1969; Lord & Novick, 1968). If relationships are relative rather than absolute, some sociological methodologists (Hargens, 1976) too argue that standardized weights should be used in seeking lawful relationships.

Table 4 uses hypothetical data to illustrate some possible effects of selection if college grades are relative rather than absolute. Two cases are shown that illustrate two different definitions of relative grading. For both definitions it is assumed that the proportions of As, Bs, Cs, etc. assigned to students do not change as a function of selection. Case 1 assumes that faculty members at more selective institutions pay attention to differences in academic potential and performance

that would be ignored at less selective colleges, so the correlation between high school and college grades also is invariant under selection. Case 2 assumes that the population correlation measures the finest distinctions that faculty members are capable of making, so the correlation changes with selection in the way shown in Table 1.

Table 4

Effects of Various Levels of Selection
on the Independent Variable of High School Grades
for Two Cases Where College Grades are Relative Rather than Absolute

	Selected Group Includes Top:			
	100%	80%	50%	20%
High School GPA				
Mean	2.60	2.84	3.15	3.50
S.D.	.70	.53	.42	.32
College GPA				
Mean	2.10	2.10	2.10	2.10
S.D.	.80	.80	.80	.80
Case 1				
Correlation Between GPAs	.65	.65	.65	.65
Regression Weights				
Standardized	.65	.65	.65	.65
Unstandardized	.74	.97	1.24	1.63
Case 2				
Correlation Between GPAs	.65	.55	.46	.36
Regression Weights				
Standardized	.65	.55	.46	.36
Unstandardized	.74	.82	.87	.91

The extensive work of experimental psychologists on adaptation-level (Helson 1948, 1959, 1964) suggests that Case 1 may describe how things work in the real world more accutately than does Case 2. When people make psychophysical judgments, the magnitude of just noticeable differences varies with changes in the dispersions of stimulus objects.

The literature on the relationship between academic potential and college performance (Astin, 1971: Chauncey & Hilton, 1965; Hoyt, 1968a, 1968b, 1968c, Munday, 1967) is more equivocal because it rarely has looked at the data in a way comparable to Table 4, but it does suggest that one or both of the cases in Table 4 describes the real world more accurately than does Table 1. While the correlations between high school and college grades vary considerably, there is little evidence for a strong systematic relationship between college selectivity and this variation. Also, average college grades increase somewhat as selectivity increases, but less than would be the case if an absolute standard of grading existed (Astin, 1971).

Several points about Table 4 should be emphasized. First, the standardized regression weights for Case 1 reveal an invariant relationship that is obscured by the unstandardized weights. Moreover, an invariant relationship of this type is just as valid for establishing scientific laws as the type of invariant relationship shown in Table 1. Second, both situations portrayed by Table 4 are genuinely different from the situation portrayed in Table 1, rather than being examples of, say, specification error. The data for Case 1 of Table 4 are quite heteroscedastic compared to the data in Table 1 and simply adding a measure of college selectivity to the data for Case 1 would not produce invariant unstandardized weights for the regression of college grades on high school grades. Similarly, correcting Case 2 of Table 4 for the effects of selection would produce constant standardized weights but varying unstandardized weights. At the same time it should be recognized that the varying unstandardized weights in Table 4 provide meaningful

information, namely that colleges differing in selectivity have different grading standards, and comparing the unstandardized weights for two colleges would provide information about such differences that would not be provided by comparing the standardized weights. In other words both the invariance revealed by the correlations and standardized weights and the differences revealed by the unstandardized weights could be scientifically important.

Another important reason for psychologists' preference for standardized weights probably is that they (1) believe most, if not all, behavioral science measures (and especially "unobserved" measures inferred from a number of indicators through factor analysis or the analysis of variance-covariance structures) are based on essentially arbitrary units of measurement, and (2) know that standardized, but not unstandardized, weights are invariant over changes in the units of measurement. Sociological methodologists explicitly recognize the latter point (Blalock, 1976; Schroeder, 1972), and in their emphasis on the desirability of interval or ratio scales could be considered the intellectual heirs of S.S. Stevens (1946). But, differences in units of measurement could still produce extremely large variation in unstandardized weights even if interval or ratio scales were achieved. Table 5 illustrates this point by showing the large changes in the unstandardized regression weights that would be produced by changing the units of measurement for a constant hypothetical relationship between the height and weight of, say, psychologists.

The bottom line of Table 5 parodies the argument that unstandardized weights are preferable because they tell you how many units of change in the dependent variable are produced by a unit of change in the

Table 5

Impact of Units of Measurement on Regression
Weights for Hypothetical Relationship Between Height and Weight

Height Measured in:	Weight Measured in:	Height		Weight		Height-Weight Correlation	Regression Weight For Predicting Weight From Height	
		Mean	S.D.	Mean	S.D.		Standardized	Unstandardized
Feet	Pounds	5.75	0.50	165.00	15.00	.60	.60	18.00
	Ounces	5.75	0.50	2,640.00	240.00	.60	.60	288.00
	Kilograms	5.75	0.50	74.84	6.80	.60	.60	8.16
	Grams	5.75	0.50	74,842.78	6,803.88	.60	.60	8,164.66
Inches	Pounds	69.00	6.00	165.00	15.00	.60	.60	1.50
	Ounces	69.00	6.00	2,640.00	240.00	.60	.60	24.00
	Kilograms	69.00	6.00	74.84	6.80	.60	.60	0.68
	Grams	69.00	6.00	74,842.78	6,803.88	.60	.60	680.39
Meters	Pounds	1.75	0.15	165.00	15.00	.60	.60	59.06
	Ounces	1.75	0.15	2,640.00	240.00	.60	.60	944.88
	Kilograms	1.75	0.15	74.84	6.80	.60	.60	26.79
	Grams	1.75	0.15	74,842.78	6,803.88	.60	.60	26,786.94
Centimeters	Pounds	175.26	15.24	165.00	15.00	.60	.60	0.59
	Ounces	175.26	15.24	2,640.00	240.00	.60	.60	9.45
	Kilograms	175.26	15.24	74.84	6.80	.60	.60	0.27
	Grams	175.26	15.24	74,842.78	6,803.88	.60	.60	267.87
Miles	Tons	.0011	.0001	.0825	.0075	.60	.60	45.00

independent variable. The unstandardized weight in this line clearly indicates that for each mile in height psychologists grow they will gain 45 tons in weight. More seriously, the unstandardized weights in this table indicate that there is no uniquely correct description of an absolute relationship between height and weight, for the description depends on the units of measurement. On the other hand the correlations and standardized regression weights clearly reveal the underlying invariance of the relationship between height and weight.

Proponents of unstandardized weights might argue (correctly) that changing the units of measurement in any given study would affect all groups to the same extent and therefore would not change the statistical significance of group comparisons. But, difficulties created by varying units of measurement seem more likely to occur when one is integrating several studies than when one is comparing groups within a single study, and different studies use different units of measurement with some frequency. Also, the psychological literature on adaptation level (Helson, 1948; 1959, 1964) suggests researchers would be likely to regard a difference between two groups as more important if that difference were 26,786.94 units than if it were 0.27 units, even if the significance levels were identical.

The closely related case of changes in the length (e.g., number of items in) a test or other behavioral science measure has received some attention from psychologists (Green, 1977; Horst, 1948), especially in their argument that test length should be considered when comparing validities (Cronbach & Glesser, 1968; Guilford, 1959). The effects of test length also are important in measurement theory because standard corrections for unreliability estimate the values that would be obtained

if the measure(s) were made infinitely long. Therefore, Table 6 illustrates the effects of changes in the length of the independent and dependent variables for hypothetical data in which all items measure exactly the same trait, have exactly the same distribution of item "difficulties" (that is, proportion of the population giving the scored response), and so on. The pattern again is clear; both the standardized and unstandardized weights change as test length changes (Green, 1977), although the unstandardized weights do vary directly with changes in the length of the dependent variable. This pattern suggests researchers should be cautious about changing the length of measures from previous studies (in order, say, to facilitate data collection) when they wish to compare results across studies.

Table 6

Impact of Changes in Test Length on Regression Weights

	Number of Items in Changed Variable			
	80	40	20	10
Unchanged Variable				
Mean	150.00	150.00	150.00	150.00
S.D.	20.00	20.00	20.00	20.00
Changed Variable				
Mean	40.00	20.00	10.00	5.00
S.D.	10.00	5.48	3.16	1.94
Correlation Between Variables	.80	.73	.63	.52
Regression Weights				
Standardized	.80	.73	.63	.52
Unstandardized for Changed:				
Independent Variable	1.60	1.43	1.26	1.04
Dependent Variable	.40	.20	.10	.05

But What About...?

There are (at least) two other important conditions that can affect regression weights but do not fit conveniently into the rubric of

comparing sociological and psychological orientations to such weights. The first condition involves selection (or other sampling variation) on the dependent variable. Any circumstance that produces missing data on the dependent variable, such as non-responses to followups in longitudinal studies, involves a form of selection, or sampling variation, on that variable. Also, researchers sometimes make a deliberate decision to select on the dependent variable. For example, sociologists studying status attainment sometimes restrict their analysis to employed persons; psychologists studying the validity of college admissions test sometimes eliminate dropouts from their analysis; etc.

This review uses colleges with open admissions policies to illustrate the effects of selection on the dependent variable. In the pure case, no selection at all on high school grades takes place in such colleges. However, students who receive low college grades are eliminated because they withdraw, are flunked out, etc., and this selection directly reduces the standard deviation of college grades and indirectly reduces both the correlation between high school and college grades and the standard deviation of high school grades. Table 7 explores the impact of various levels of selection on the dependent variable of college grades. This analysis, of course, again assumes an absolute relationship between the independent and dependent variables.

The pattern of results for the correlations and standardized regression weights is identical with the pattern in Table 1, but the pattern of results for the unstandardized regression weights is quite different in that these weights were strongly affected by selection

Table 7

Effects of Various Levels of Selection on the Dependent Variable of College Grades

	Surviving Group Includes Top:			
	100%	80%	50%	20%
High School GPA				
Mean	2.60	2.76	2.97	3.24
S.D.	.70	.64	.60	.57
College GPA				
Mean	2.10	2.37	2.74	3.22
S.D.	.80	.61	.48	.36
Correlation Between GPAs	.65	.55	.46	.36
Regression Weights				
Standardized	.65	.55	.46	.36
Unstandardized	.74	.53	.37	.23
Population/Surviving Group Regression Weight Ratios				
Standardized	1.00	1.19	1.42	1.79
Unstandardized	1.00	1.41	2.03	3.20

on the dependent variable. Indeed, as the last two rows in this table indicate, the unstandardized weights were affected more than the standardized weights. Contrary to what some methodologists (Schoenberg, 1972) have suggested, therefore, a finding that the unstandardized weights for two groups differ more than the standardized weights may be substantively unimportant, indicating only that the two groups differ in degree of selection on the dependent variable. The reasons for the differences between Table 1 and Table 7 are easy to understand if one remembers that the relationship between two variables includes two regressions, and realizes that what is invariant under selection is the regression of the incidental selection variable on the explicit selection variable. Therefore, the regression of the independent variable on the dependent variable is invariant under selection on the dependent variable. Unfortunately, this is the wrong regression for

making causal inferences from unstandardized weights; such inferences necessarily involve the regression of the dependent variable on the independent variable.

The implications of the results in Table 7 could be quite important, so it appeared desirable to explore their generality in a situation where the causal relationships are clearer. Accordingly artificial data were constructed for the Kim and Mueller (1976) example shown in Table 2. To permit exact replication of the analysis reported here, the actual artificial data are shown in Table 8. These artificial data consider the psychometrically ideal case of 50%-50% splits on sex and race and no correlation among sex, race and experience, and also assume that workers have 0 to 6 years of experience. Given these assumptions, data generated by formula (1) are identical with data generated by the formula:

$$\text{Pay Rate} = \$13 + \$2 \times \text{Sex} + \$3 \times \text{Race} + E \qquad (2)$$

where E is random measurement error with Mean=0 and Standard Deviation=1 ($).

These data were used to compute the consequences of different levels of selection on the dependent variable of pay rate. In a real study such selection could occur if, say, workers below the salary level required for participation in a pension plan or for receiving a permanent contract were eliminated. The impact of selection was computed for three cases. Case I treated all three of the variables of sex, race, and experience as causal variables. In terms of the original example, this case was correctly specified and error free. Case II treated only sex and race as causal variables and therefore examined the

Table 8

Artificial Data For Kim and Muller's Example

Case 1 Years of Experience	Case 2 Random Measurement Error		Black Females	Black Males	White Females	White Males
6	+3	N	1	1	1	1
		Pay Rate	$16	$18	$19	$21
5	+2	N	4	4	4	4
		Pay Rate	$15	$17	$18	$20
4	+1	N	25	25	25	25
		Pay Rate	$14	$16	$17	$19
3	0	N	40	40	40	40
		Pay Rate	$13	$15	$16	$18
2	-1	N	25	25	25	25
		Pay Rate	$12	$14	$15	$17
1	-2	N	4	4	4	4
		Pay Rate	$11	$13	$14	$16
0	-3	N	1	1	1	1
		Pay Rate	$10	$12	$13	$15

effects of selection on an analysis that was subject either to specification error (if the additional variance was determined by years of experience) or to measurement error on the dependent variable (if the additional variance was random). Case III, a case that can be considered another illustration of specification error, examined separately the bivariate relationships of sex, race, and experience with pay rate. The results for all three cases are presented in Table 9.

The pattern revealed by this table is complex. The results provide some additional support for sociologists' preference for unstandardized weights in that the unstandardized weights were close to the underlying parameters so long as the analysis was correctly specified, the dependent variable was error-free, and selection was not extreme. The standardized weights actually increased under selection in this example because

Table 9

Effect of Selection on Dependent Variable of Pay Rate for Artificial Data from Kim and Muller's Example

	Retained Group Includes Top:			
	100%	81%	50%	19%
Case I: Correctly Specified, Error Free Condition				
Standardized Weight For:				
Sex	.49	.61	.81	1.25
Race	.73	.91	.95	.52
Experience	.49	.61	.84	.75
Unstandardized Weight For:				
Sex	2.00	1.99	1.92	3.31
Race	3.00	3.00	3.00	3.00
Experience	1.00	1.00	1.05	.61
Case II: Condition of Specification or Measurement Error				
Standardized Weight For:				
Sex	.49	.50	.62	.12
Race	.73	.77	.56	.09
Unstandardized Weight For:				
Sex	2.00	1.64	1.49	.31
Race	3.00	2.55	1.78	.51
Case III: Bivariate Condition				
Standardized Weight For:				
Sex	.49	.30	.46	.12
Race	.73	.65	.38	.09
Experience	.49	.37	.31	.70
Unstandardized Weight For:				
Sex	2.00	.99	1.10	.31
Race	3.00	2.12	1.21	.49
Experience	1.00	.61	.39	.56

selection on the dependent variable produced increasingly large negative correlations among sex, race, and experience, and these correlations played an increasing role in determining the standardized weights.

Because dependent variables in the behavioral sciences rarely if ever are completely free of measurement error, the most relevant results in Table 9 are those for Case II, which examines selection in the presence of specification error or measurement error on the dependent

variable. When either of these errors was combined with selection or sampling variation on the dependent variable, both the standardized and unstandardized weights were strongly affected. (Those results contrast with those in Table 3 which show that in the absence of selection on the dependent variable the unstandardized weights are unaffected by dependent variable measurement error.) Similarly, Case III confirms that when bivariate relationships are subject to selection on the dependent variables, both regression weights are strongly affected.

One special feature of the data in Table 9 should be mentioned. The computed changes in the regression weights are greater than the theoretical changes predicted by Gulliksen's (1950) formulas (Richards, Note 1). The most likely explanation for these outcomes is that the regressions of sex and race on the explicit selection variable of pay rate were heteroscedastic, non-linear, or both (Greener and Osburn, 1979). For example, the reversals and "impossible" standardized weights in the 50% and 19% groups occurred because much of the variance on sex and race had been eliminated in those groups. Similar effects could occur whenever dummy variables are included in an analysis that combines specification or measurement error with selection on the dependent variable, and would be more serious if the splits on the dichotomous variables were more extreme.

The final condition to be discussed in this paper involves explicit selection on a mediating variable. This type of selection may occur fairly frequently in behavioral science research. For example, researchers studying contraceptive behavior sometimes limit their analyses to married persons; researchers studying the impact of schooling sometimes restrict

their analyses to students in particular "tracks," or at best compare students assigned non-randomly to different tracks; etc.

The academic selection example used in Tables 1 and 7 can be extended to this type of selection by considering the additional variable of college graduation. The relationship between high school grades and college graduation is mediated through college grades. Therefore, in colleges with open admissions policies the elimination of students with low college grades constitutes explicit selection on a mediating variable. If we place these relationships in a path analytic context, they can be described by the path diagram shown in Figure 1. Residual, or disturbance, variables are irrelevant to the present discussion, so such variables are not shown on this diagram.

Hypothetical data for college graduation were added to the academic selection example used previously. The mean and standard deviation for college graduation were chosen to reflect findings that among those (whites) who go to college about half graduate (Golladay and Nevell, 1978). Then the relationships among high school grades, college grades, and college graduation were analyzed for various levels of selection on college grades, following the path diagram in Figure 1, and using the procedures described by Alwin and Hauser (1975) to decompose the relationship between high school grades and college graduation into total, direct, and indirect components. Results are shown in Table 10.

This pattern too is fairly complex. The standardized weights, or path coefficients, for the relationship between college grades and graduation change with varying degrees of selection while the unstandardized weights, or path regression coefficients, are invariant. The unstan-

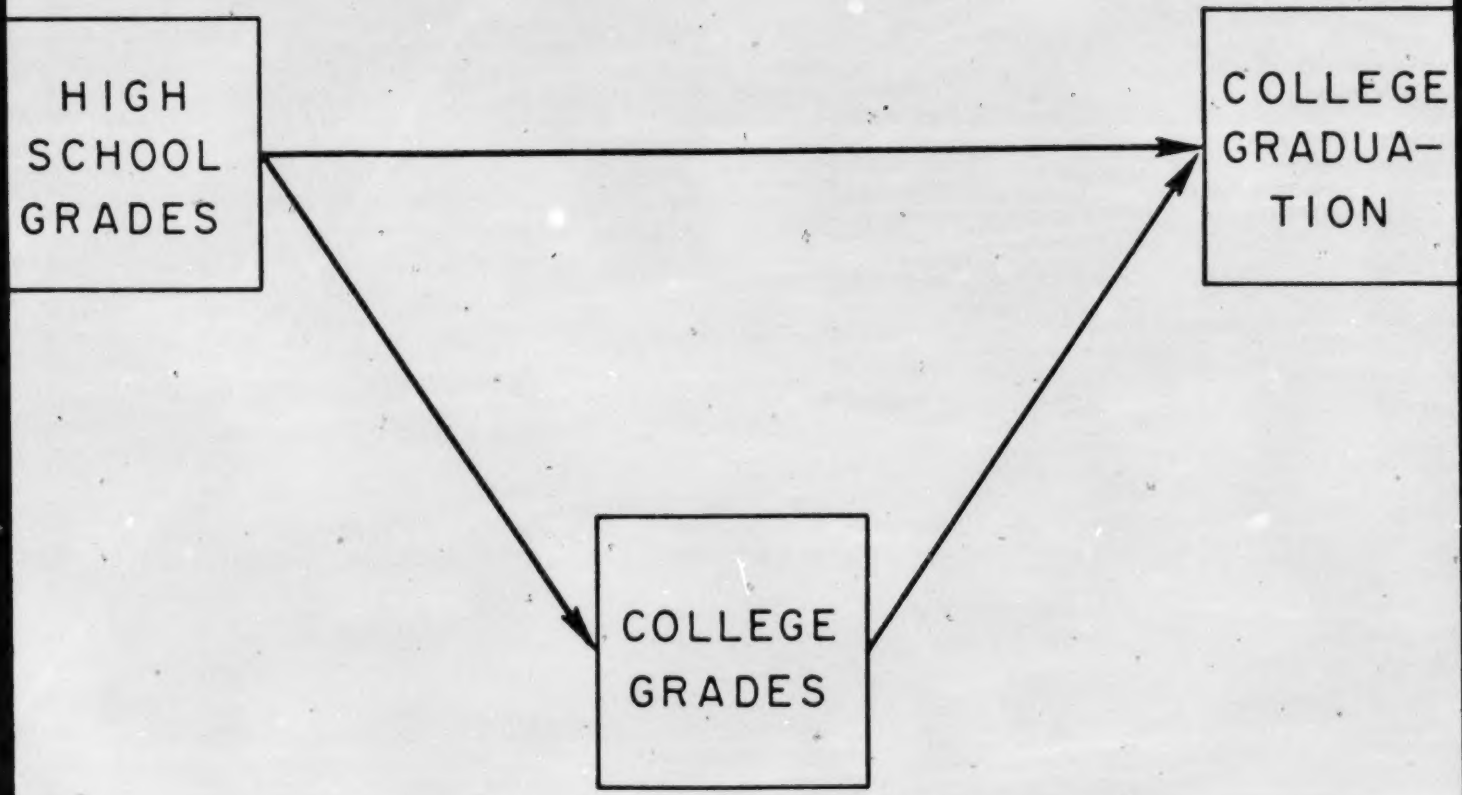

Figure 1. Path diagram of relationships among high school grades, college grades, and college graduation.

Table 10

Path and Path Regression Coefficients for Various Levels of Selection on the Mediating Variable of College Grades

	Hypothetical Relationships Among Variables in Population				
	Intercorrelations				
	1	2	3	Mean	S.D.
1. High School GPA	--			2.60	0.70
2. College GPA	.65	--		2.10	0.80
3. College Graduation	.60	.70	--	0.50	0.50

	Surviving Group Includes Top			
	100%	80%	50%	20%
Standardized or Path Coefficients				
Influence of High School GPA				
Total	.60	.51	.44	.37
Direct	.25	.26	.26	.26
Indirect	.35	.25	.18	.11
Influence of College GPA	.54	.46	.38	.31
Unstandardized or Path Regression Coefficients				
Influence of High School GPA				
Total	.43	.36	.30	.26
Direct	.18	.18	.18	.18
Indirect	.25	.18	.12	.08
Influence of College GPA	.34	.34	.34	.34

dardized weights for the direct effect of high school grades also are invariant, and the standardized weights change very little, actually increasing slightly. Most important, both the standardized and unstandardized weights for the total and indirect effects of high school grades are strongly affected by explicit selection on the mediating variable of college grades, with the unstandardized weights being somewhat more affected than the standardized weights. (This can be verified by computing ratios similar to those shown in Table 7.) The ratio of indirect to

direct effects also would be strongly affected. These results militate against the kinds of inferences proponents of unstandardized weights wish to make from them.

Implications

The intent of this review was to further communication among researchers by clarifying the standardized vs. unstandardized weights issue. The most basic difference in perspectives on regression weights seems to involve different beliefs about the nature of the universe and of behavioral science data. Those researchers who believe that behavioral phenomena involve absolute relationships between universal qualities are likely to prefer unstandardized weights because such weights are somewhat more invariant when relationships are absolute. On the other hand, those researchers who believe that behavioral phenomena usually involve relative relationships are likely to prefer standardized weights because such weights are somewhat more invariant when relationships are relative. To some degree, these differing beliefs should be amenable to empirical resolution. For example, it would be relatively easy to investigate which of Table 1, Case 1 of Table 4, or Case 2 of Table 4 best describes the impact of college selectivity on the relationship between high school grades (or admission test scores) and college grades. Both the American College Testing Program and the College Entrance Examination Board have obtained extensive data through their respective validity study services that would permit such an investigation. Whatever the outcome of such an investigation, however, the results would apply only to one specific relationship, and a series of studies of other specific relationships might well find that some relationships are relative, some are absolute,

and a relatively large number are neither clearly relative nor clearly absolute.

In short, there probably is no one, unequivocal answer to the question of whether standardized or unstandardized regression weights are preferable. The illustrations presented in this review indicate that each kind of weight is invariant in some circumstances where the other kind is not, and each kind can provide scientifically meaningful information that the other kind cannot. One conclusion to be drawn from these illustrations, therefore, is that psychologists might well emulate the increasingly prevalent practice among sociologists of reporting both weights. More importantly, these illustrations imply that sociologists' enthusiasm for unstandardized weights is excessive, while psychologists' preference for standardized weights is oversimplified.

In the same vein, those illustrations raise doubts about the usefulness of all regression techniques in non-randomized, quasi-experimental contexts. A number of circumstances that could affect regression weights were examined, and in most cases these circumstances affect both the standardized and the unstandardized weights, sometimes quite strongly. Thus neither kind of weight appears particularly robust, and certainly neither kind appears to provide a panacea for overcoming the problems of analyzing behavioral data. Presumably these and other similar considerations underlie the scepticism expressed by other writers (Campbell & Erlebacher, 1975; Cook & Campbell, 1976, pp. 254-288; Gilbert, Mosteller & Tukey, 1976) about the appropriateness of any regression technique for non-experimental causal analysis.

Reference Notes

1. Richards, J.M., Jr. Standardized vs. unstandardized regression weights: Once more into the breach. Manuscript submitted for publication, 1979.

References

Alwin, D. F. & Hauser, R. M. The decomposition of effects in path analysis. American Sociological Review, 1975, 40 37-42.

American College Testing Program. College student profiles. Iowa City, IA: Author, 1966.

Astin, A. W. Predicting academic performance in college: Selectivity data for 2,300 American colleges. New York, NY: Free Press, 1971.

Blalock, H. M., Jr. Causal inferences in nonexperimental research. Chapel Hill, NC: University of North Carolina Press, 1964.

Blalock, H. M., Jr. Path coefficients versus regression coefficients. American Sociological Review, 1967, 32, 675-676.

Blalock, H. M., Jr. The measurement problem: A gap between the language of theory and research. In H. M. Blalock, Jr. and A. B. Blalock (Eds.) Methodology in social research. San Francisco, CA: Jossey-Bass Publishing Co., 1968.

Blalock, H. M., Jr. Can we find a genuine ordinal slope analogue? In D. R. Heise (Ed.) Sociological methodology 1976. San Francisco, CA: Jossey-Bass Publishing Co., 1976.

Bohrnstedt, G. W. Observations on the measurement of change. In E. Borgatta (Ed.) Sociological methodology 1969. San Francisco, CA: Jossey-Bass Publishing Co., 1969.

Bohrnstedt, G. W. & Carter, T. M. Robustness in regression analysis. In H. L. Costner (Ed.) Sociological methodology 1971. San Francisco, CA: Jossey-Bass Publishing Co., 1971.

Campbell, D. T. & Erlebacher, A. How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education

look harmful. In E. L. Struening and M. Guttentag (Eds.) Handbook of evaluation research. Beverly Hills, CA: Sage Publications, 1975.

Chauncey, H. & Hilton, T. L. Are aptitude tests valid for the highly able? Science, 1965, 148, 1297-1304.

Cook, T. D. & Campbell, D. T. The design and conduct of quasi-experiments and true experiments in field settings. In M. D. Dunnette (Ed.) Handbook of industrial and organizational psychology. Chicago, Ill.: Rand McNally College Publishing Co., 1976.

Cooley, W. W. Explanatory observational studies. Educational Researcher, 1978, 7, 9-15.

Cronbach, L. J. & Gleser, G. Psychological tests and personnel decisions. Urbana, Ill.: University of Illinois Press, 1968.

Dawes, R. M. & Corrigan, B. Linear models in decision making. Psychological Bulletin, 1974, 81, 95-106.

Duncan, O. D. Path analysis: Sociological examples. American Journal of Sociology, 1966, 72, 1-16.

Duncan, O. D. Introduction to structural equation models. New York, NY: Academic Press, 1975.

Galloway, M. A. & Noell, J. The condition of education. Washington, D.C.: National Center for Educational Statistics, 1978.

Ghiselli, E. E. Theory of psychological measurement. New York, NY: McGraw-Hill Book Co., 1964.

Gilbert, J. P., Mostellar, F., & Tukey, J. Steady social progress requires quantitative evaluation to be searching. In C. C. Abt (Ed.) The evaluation of social programs. Beverly Hills, CA: Sage Publications, 1976.

Green, B. F. Parameter sensitivity in multivariate methods. Multivariate Behavioral Research, 1977, 12, 263-287.

Greener, J. M. & Osburn, H. G. An empirical study of the accuracy of corrections for restriction of range due to explicit selection. Applied Psychological Measurement, 1979, 3, 31-41.

Guilford, J. P. Psychometric methods. New York, NY: McGraw-Hill Book Co., 1954.

Gulliksen, H. Theory of mental tests. New York, NY: John Wiley & Sons, 1950.

Hargens, L. L. A note on standardized coefficients as structural parameters. Sociological Methods & Research, 1976, 5, 247-256.

Helson, H. Adaptation-level as a basis for a quantitative theory of frames of reference. Psychological Review, 1948, 55, 297-313.

Helson, H. Adaptation-level theory. In S. Koch (Ed.) Psychology: A Study of a science. New York, NY: McGraw-Hill Book Co., 1959.

Helson, H. Adaptation-level theory: An experimental and systematic approach to behavior. New York, NY: Harper & Row, 1964.

Horst, P. Regression weights as a function of test length. Psychometrika, 1948, 13, 125-134.

Hoyt, D. P. Description and prediction of diversity among four-year colleges. Measurement and Evaluation in Guidance, 1968, 1, 16-26. (a)

Hoyt, D. P. Description and prediction of diversity among two-year colleges. Personnel and Guidance Journal, 1968, 46, 997-1004. (b)

Hoyt, D. P. Generalized academic prediction in four-year colleges. Personnel and Guidance Journal, 1968, 47, 130-136. (c)

Jöreskog, K. G. A general method for estimating a linear structural equation system. In A. S. Goldberger and O. D. Duncan (Eds.) Structural equation models in the social sciences. New York, NY: Seminar Press, 1973.

Kim, J. & Mueller, C. W. Standardized and unstandardized coefficients in causal analysis. Sociological Methods & Research, 1976, 4, 423-437.

Laughlin, J. E. Comment on "Estimating coefficients in linear models: It don't make no nevermind." Psychological Bulletin, 1978, 85, 247-253.

Lavin, D. E. The prediction of academic performance: A theoretical analysis and review of research. New York, NY: Russell Sage Foundation, 1965.

Linn, R. L. Range restriction problems in the use of self-selection groups for test validation. Psychological Bulletin, 1968, 69, 69-73.

Lord, F. M. & Novick, M. R. Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley, 1968.

Munday, L. Predicting college grades from ACT data. Educational and Psychological Measurement, 1967, 27, 401-406.

Nie, N. H., Hull, C. H,, Jenkins, J. G., Steinbrenner, K., & Bent, D. H. Statistical package for the social sciences. (2nd ed.) New York, NY: McGraw-Hill Book Co., 1975.

Pruzek, R. M. & Frederick, B. C. Weighting predictors in linear models: Alternatives to least squares and limitations of equal weights. Psychological Bulletin, 1978, 85, 254-266.

Schoenberg, R. Strategies for meaningful comparisons. In H.C. Costner (Ed.) Sociological Methodology 1972. San Francisco, CA: Jossey-Bass Publishing Co., 1972.

Stevens, S. S. On the theory of scales of measurement. Science, 1946, 103, 670-680.

Wainer, H. Estimating coefficients in linear models: It don't make no nevermind. Psychological Bulletin, 1976, 83, 213-217.

Wainer, H. On the sensitivity of regression and regressors. Psychological Bulletin, 1978, 85, 254-266.

Werts, C. E. & Linn, R. L. Path analysis: Psychological examples. Psychological Bulletin, 1970, 74, 193-212.